

# The Two Envelope Problem: a Paradox or Fallacious Reasoning?

Aris Spanos

Department of Economics,  
Virginia Tech, Blacksburg, VA 24061  
<aris@vt.edu>

January 2013

## Abstract

The primary objective of this note is to revisit the two envelope problem and propose a simple resolution. It is argued that the paradox arises from the ambiguity associated with the money content  $\$x$  of the chosen envelope. When  $X=x$  is observed it is not known which one of the two events,  $X=\theta$  or  $X=2\theta$ , has occurred. Moreover, the money in the other envelope  $Y$  is not independent of  $X$ ; when one contains  $\theta$  the other contains  $2\theta$ . By taking these important features of the problem into account, the paradox disappears.

## 1 Introduction

Consider two *indistinguishable* envelopes that contain  $\$ \theta > 0$  and  $\$ 2\theta$ . Player 1 chooses one of the envelopes at random and observes its content  $X=x$ <sup>1</sup>. The player is given the choice to either keep  $\$x$ , or exchange it with the contents  $Y$  of the other envelope. What should player 1 do?

The traditional account is that ‘rational’ reasoning by player 1 will evaluate the expected value of  $Y$ , defined in terms of  $x$ :

$$\begin{aligned} Y &= \frac{x}{2}, & \mathbb{P}(Y = \frac{x}{2}) &= \frac{1}{2}, \\ Y &= 2x, & \mathbb{P}(Y = 2x) &= \frac{1}{2}. \end{aligned}$$

Hence, player’s 1 expected winnings by trading envelopes will be:

$$E(Y) = \left(\frac{x}{2}\right) \mathbb{P}(Y = \frac{x}{2}) + 2x \mathbb{P}(Y = 2x) = \left(\frac{x}{2}\right) \left(\frac{1}{2}\right) + 2x \left(\frac{1}{2}\right) = \frac{5}{4}x > x. \quad (1)$$

This suggests that it will be rational for player 1 *to always exchange his envelope*, whatever the value  $x$ . This reasoning is clearly fallacious, but the difficulty is

---

<sup>1</sup>In certain variants of the paradox the player does not see  $x$ , but that makes no difference to the following discussion.

to pinpoint the source of the problem; see Nalebuff (1989), Broome (1995), Chalmers (2002), Clark and Shackel (2000), Dietrich and List (2004), Falk and Nickerson (2009) inter alia.

It is argued that the paradox arises because of the ambiguity of the event  $X=x$  stemming from the fact that the observed value  $x$  stands for two different but *unknown* values  $\theta$  or  $2\theta$ . That is, when  $X=x$  is observed one does *not* know which event  $X=\theta$  or  $X=2\theta$  has occurred. Moreover, the traditional account uses the marginal distribution of  $Y$  expressed in terms of the equivocal event  $X=x$ , when in fact the random variables  $X$  and  $Y$  are dependent; when one envelope contains  $\theta$  the other contains  $2\theta$ . When these features are taken into account the paradox vanishes.

## 2 A paradox or fallacious reasoning?

The first issue to reconsider is the nature of the random variable  $X$  denoting the money in the envelope initially chosen by Player 1. The player observes its content  $X=x$ , but does not know is whether  $x$  represents  $\theta$  or  $2\theta$ . Hence, the random variable  $X$  is, in effect, *latent*:

$$X = \begin{cases} \theta & \text{for } x=\theta \\ 2\theta & \text{for } x=2\theta \end{cases}, \quad \begin{matrix} \mathbb{P}(X=\theta)=.5 \\ \mathbb{P}(X=2\theta)=.5 \end{matrix}$$

The probability .5 arises from the fact that the two envelopes are indistinguishable. In light of the fact that when one of the envelopes contains  $\$ \theta$  the other must contain  $\$ 2\theta$ , the relevant distribution is the joint distribution of  $X$  and  $Y$ , given in table 1. It is important to note is that the support of both random variables,  $R_X=\{x: f(x) > 0\}$  and  $R_Y=\{y: f(y) > 0\}$ , depends on the unknown parameter  $\theta$ , rendering them *non-regular*; see Cox and Hinkley (1974).

Table 1			
$X \setminus Y$	$\theta$	$2\theta$	$f(x)$
$\theta$	0	.5	.5
$2\theta$	.5	0	.5
$f(y)$	.5	.5	1

(2)

Not surprisingly,  $f(x, y) \neq f(x) \cdot f(y)$ , for all  $(x, y)$ , and thus  $X$  and  $Y$  are *not* independent, but the problem is now symmetric with respect to both random variables. Indeed, the expected winnings from either envelope are identical:

$$\begin{aligned} E(X) &= .5\theta + .5(2\theta) = 1.5\theta, \\ E(Y) &= .5\theta + .5(2\theta) = 1.5\theta, \end{aligned} \tag{3}$$

rendering player 1 indifferent between retaining  $\$ x$  or exchanging envelopes. This result shows that whether player 1 should exchange depends crucially on the relationship between the observed value  $x$  and  $\theta$ :

- (i) for  $x=2\theta$ ,  $E(Y)=\frac{3}{4}x < x$ , and thus player 1 should *not* exchange, but
- (ii) for  $x=\theta$ ,  $E(Y)=\frac{3}{2}x > x$ , and player 1 should exchange.

The problem, however, is that observing  $X=x$  is inadequate to make an informed decision whether to exchange or not. This resolves the paradox!

### 3 Conditioning on latent variables

A more circuitous but illuminating way to reach the same conclusion is to treat both random variables as latent and deal with the ambiguity of the event  $X=x$  using the conditional expectation  $E(Y|\sigma(X))$ , where  $\sigma(X)=\{S, \emptyset, X=\theta, X=2\theta\}$  denotes the sigma-field generated by  $X$ . Since  $\sigma(X) \subset \mathcal{F}$ , conditioning on  $\sigma(X)$  simply acknowledges the possible events generated by  $X$  via restricting the universal  $\mathcal{F}$  related to the original probability space  $(S, \mathcal{F}, \mathbb{P}(\cdot))$ , upon which both random variables  $(X, Y)$  have been defined. Formally, conditioning on  $\sigma(X)$  constitutes a restriction because:

$$E(Y|\mathcal{F})=Y \text{ but } E(Y|\sigma(X))=g(X) \neq Y.$$

Moreover, the random variable  $E(Y|\sigma(X))$  does not depend on the *particular values*  $x$  of  $X$  because for any Borel function  $h(\cdot)$  which keeps those values distinct, i.e. for two different values of  $X$ , say  $x_1 \neq x_2$ ,  $h(x_1) \neq h(x_2)$  (Renyi, 1970, p. 259):

$$E(Y|\sigma(X))=E(Y|\sigma(h(X))), \text{ since } \sigma(X)=\sigma(h(X)).$$

To evaluate  $E(Y|\sigma(X))$  one needs both conditional distributions:

$$f(Y|X=\theta)=\begin{cases} \frac{f(y=2\theta, x=\theta)}{f(x=\theta)}=\frac{.5}{.5}=1, & \text{for } Y=2\theta \\ \frac{f(y=\theta, x=\theta)}{f(x=\theta)}=\frac{0}{.5}=0, & \text{for } Y=\theta \end{cases}$$

$$f(Y|X=2\theta)=\begin{cases} \frac{f(y=2\theta, x=2\theta)}{f(x=2\theta)}=\frac{0}{.5}=0, & \text{for } Y=2\theta \\ \frac{f(y=\theta, x=2\theta)}{f(x=2\theta)}=\frac{.5}{.5}=1, & \text{for } Y=\theta \end{cases}$$

Hence,  $E(Y|\sigma(X))$  defines a random variable of the form:

$$E(Y|\sigma(X))=[2\theta + 0 \cdot \theta] \mathbb{I}_{\{x=\theta\}} + [\theta + 0 \cdot 2\theta] \mathbb{I}_{\{x=2\theta\}} = 2\theta \mathbb{I}_{\{x=\theta\}} + \theta \mathbb{I}_{\{x=2\theta\}}, \quad (4)$$

where  $\mathbb{I}_{\{x=\theta\}}$  is the indicator function. To derive the expected winnings of exchanging envelopes one needs  $E(Y)$  which can be derived from (4) using the law iterated expectations (Williams, 1991):

$$E(Y)=E_X\{E(Y|\sigma(X))\} = 2\theta(.5) + \theta(.5) = 1.5\theta, \quad (5)$$

which coincides with the result in (3).

## 4 The fallacy and the induced distribution of $\theta$

One might object to the reasoning giving rise to the evaluation of  $E(Y|\sigma(X))$  in (5) by claiming that one *can* attach probabilities to  $x=\theta$  and  $x=2\theta$ . Indeed, this has been the basis of several Bayesian solutions to this paradox that often revolve around the conditional probabilities:

$$\mathbb{P}(X=x|\theta=x), \quad \mathbb{P}(X=x|\theta=\frac{x}{2}),$$

stemming from some form of prior information; see Christensen and Utts (1992) and Lindley (2006). This move, however, invokes the potential ambiguity between the event  $X=\theta$  and the value assignment  $\theta=x$ ; the former is a legitimate frequentist event, but the latter constitutes an event only in the context of Bayesian inference. This ambiguity can inadvertently give rise to creating an induced distribution for  $\theta$ . This can easily arise when an overlap between the parameter and sample spaces has been created by a non-regular distribution. This overlap could misleadingly be used to derive the ‘induced’ distribution of  $\theta$  from that of  $X$  :

$$\begin{array}{|c|c|} \hline \textbf{Table 2} \\ \hline x & f(x) \\ \hline \theta & .5 \\ \hline 2\theta & .5 \\ \hline \end{array} \Rightarrow \begin{array}{|c|c|} \hline \textbf{Table 3} \\ \hline \theta & p(\theta) \\ \hline x & .5 \\ \hline \frac{x}{2} & .5 \\ \hline \end{array} \quad (6)$$

and then (inadvertently) proceed to use  $p(\theta)$  in place of  $f(x)$ .

To demonstrate how conflating  $X=\theta$  and  $X=2\theta$  with  $x=\theta$  and  $x=2\theta$  can lead to the fallacious result (1), consider replacing  $f(x)$  (table 2) with  $p(\theta)$  (table 3) in (4). This replacement yields:

$$E^\dagger(Y|\sigma(X)) = 2x\mathbb{I}_{\{x=\theta\}} + \frac{x}{2}\mathbb{I}_{\{x=2\theta\}} \Rightarrow \quad (7)$$

$$E^\dagger(Y) = E_x\{E^\dagger(Y|\sigma(X))\} = 2x(\frac{1}{2}) + \frac{x}{2}(\frac{1}{2}) = \frac{5}{4}x,$$

which coincides with the fallacious expected value in (1).

This confusion can be seen in the evaluation of the likelihood function:

$$\begin{aligned} L(\theta=x) &= \mathbb{P}(X=x|\theta=x) = \mathbb{P}(X=\theta|\theta=x) = .5, \\ L(\theta=\frac{x}{2}) &= \mathbb{P}(X=x|\theta=\frac{x}{2}) = \mathbb{P}(X=2\theta|\theta=\frac{x}{2}) = .5, \end{aligned} \quad (8)$$

given in Pawitan (2001), p. 27.

## 5 Conclusion

The key conclusion is that the two envelope (exchange) paradox stems primarily from the ambiguity associated with the money content  $\$x$  of the chosen envelope. When the event  $X=x$  is observed one does not know which of the two different events,  $X=\theta$  or  $X=2\theta$ , has occurred. Moreover, the money content of the other envelope  $Y$  is dependent on  $X$ ; if one contains  $\$ \theta$  the other contains  $\$ 2\theta$ . The

appropriate way to deal with these features of the problem is to use treat both random variables as *latent* and derive  $E(Y)$  either directly or via  $E(Y|\sigma(X))$ .

Taking these features into account resolves the paradox because:

$$E(Y)=E_X\{E(Y|\sigma(X))\}=1.5\theta \neq E^\dagger(Y)=\left(\frac{x}{2}\right)\mathbb{P}\left(Y=\frac{x}{2}\right)+2x\mathbb{P}(Y=2x)=1.25x.$$

The result  $E(Y)=1.5\theta$  indicates that the optimal strategy for player 1 depends crucially on whether  $x=\theta$  or  $x=2\theta$ . Without the latter information, player 1 is indifferent between the two envelopes since  $E(Y)=E(X)=1.5\theta$ .

## References

- [1] Broome, J. (1995), “The two envelope paradox,” *Analysis*, **55**: 6-11.
- [2] Chalmers, D. J. (2002), “The St. Petersburg two-envelope paradox,” *Analysis*, **62**: 155-157.
- [3] Christensen, R. and J. Utts (1992), “Bayesian Resolution of the ‘Exchange Paradox’,” *The American Statistician*, **46**: 274-276.
- [4] Clark, M., and Shackel, N. (2000), “The two envelope paradox,” *Mind*, **109**: 415-442.
- [5] Cox, D. R. and D. V. Hinkley (1974), *Theoretical Statistics*, Chapman & Hall, London.
- [6] Dietrich, F., and List, C. (2004), “The two-envelope paradox: An axiomatic approach,” *Mind*, **114**: 239-248.
- [7] Falk, R. and R. Nickerson. (2009), “An inside look at the two envelopes paradox,” *Teaching Statistics*, **31**: 39-41.
- [8] Lindley, D. V. (2006), *Understanding Uncertainty*, Wiley, NY.
- [9] Nalebuff, B. (1989), “Puzzles: the other person’s envelope is always greener,” *Journal of Economic Perspectives*, **3**: 171-181.
- [10] Pawitan, Y. (2001), *In All Likelihood: Statistical Modeling and Inference Using the Likelihood*, Oxford University Press, Oxford.
- [11] Renyi, A. (1970), *Foundations of Probability*, Holden-Day, San Francisco.
- [12] Williams, D. (1991), *Probability with Martingales*, Cambridge University Press, Cambridge.